Time Based Version Management with FairDB

Denis Bertini[1]
[1]Scientific Computing ,GSI, Darmstadt, Germany

Introduction

Because of unprecedented interaction rate (10^7/s) and data rates from 100Gb/s up to 1Tb/s, experiments at Fair [1][2] employ self-triggered front-end electronics featuring a continuously sampling system where the detector subsystems are synchronized by a precision time stamp distribution system. In this context of free data streaming, time or more precisely time interval is the natural detector buffer identifier.
For this reason, FairDB [3] implements internally a version management for the parameter fully based on time.

Temporal Database

Conceptually in a relational database data is organized in two-dimensional space (attribute, tuple). FairDB implements a time-based version management by adding the time as a third dimension leading to a three dimensional model (attribute, tuple, time) as shown in Figure [1].

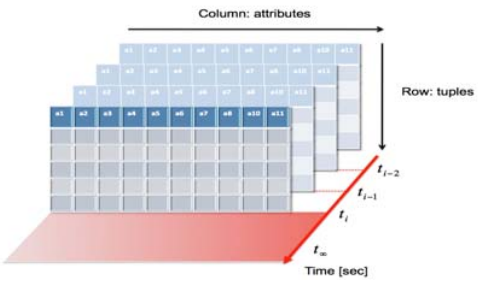

Figure 1: FairDB timed based versioning uses time as additional dimension to the relational data model.

Relational Model

FairDB separates the metadata table (logical data) from the payload data table (physical data). It uses a two-level hierarchy to store the physical data which is illustrated in Figure [2]. This two-level hierarchy data access corresponds to a two-table structure for each type of parameter data in the FairDB relational model. The two table are connected via a unique identifier primary key used as a sequence number.

Validity Range

Ultimately, any of the data retrieved could depend on the run or the time interval defining the event being processed. Detector relevant parameters, such as calibration constants,

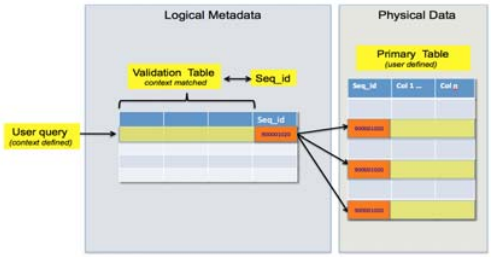

Figure 2: Data Table location Hierarchy

will change with time and the interface has to retrieve the correct ones for the current event time-slice. For this reason, all requests for data through the interface must supply information about:

- The date and time interval of event (in UTC)
- The type of data: real or Monte Carlo
- The type of Detector.

In FairDB this information is called a Context. In the database all information is tagged by a Context i.e by a validation range which identifies the type of data and detector and the ranges of date times for which it is valid.

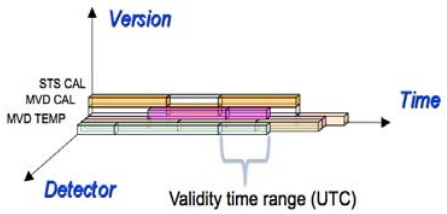

Figure 3: FairDB Multidimensional version management. Only the Validity Time Range is mandatory for parameter retrieval.

References

[1] The Data Acquisition and Trigger System of Panda, 2008 IEEE N57-1

[2] Simulation and reconstruction of free-streaming data in CBM, 2011 J. Phys.: Conf. Ser. 331 032008

[3] FairRoot Virtual Database (User Manual). `https://panda-wiki.gsi.de/foswiki/pub/Computing/PandaRoot/FairRootVirtualDatabase.pdf`